Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Incorporation of a disaccharide nucleoside into the backbone of double-stranded DNA: crystallization and preliminary X-ray diffraction

Kristof Van Hecke,[a] Koen Uytterhoeven,[a] Sergey N. Mikhailov,[b] Piet Herdewijn[c] and Luc Van Meervelt[a]*

[a]Biomolecular Architecture, Department of Chemistry, Katholieke Universiteit Leuven, Celestijnenlaan 200 F, B-3001 Leuven (Heverlee), Belgium, [b]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov Street 32, Moscow 119991, Russia, and [c]Laboratory for Medicinal Chemistry, Rega Institute, Katholieke Universiteit Leuven, Minderbroederstraat 10, B-3000 Leuven, Belgium

Correspondence e-mail: luc.vanmeervelt@chem.kuleuven.be

Received 13 September 2005
Accepted 22 September 2005
Online 30 September 2005

Incorporation of a disaccharide nucleoside into double-stranded DNA can be considered as a chemical (non-enzymatic) alternative for site-specific cleavage of DNA. Crystals of the sequence d(CGCGAATT*CGCG), where * is an incorporated ribose, were obtained by hanging-drop vapour diffusion and diffracted to 2.6 Å. The crystals belong to the orthorhombic space group $P222_1$, with unit-cell parameters a = 41.52, b = 57.63, c = 81.39 Å, indicating a new crystal packing motif for an oligonucleotide dodecamer sequence.

1. Introduction

Systematic cleavage of double-stranded DNA at specific sites requires the use of restriction enzymes. Most restriction enzymes recognize specific sequences of four to eight bases and produce dsDNA fragments with either blunt or sticky ends. When produced by the same restriction enzyme, cohesive fragments can always be stuck together and joined permanently by DNA ligase. This technique has lead to the development of numerous applications in recombinant DNA technology, *e.g. in vitro* site-specific mutagenesis.

Alternatively, the use of an incorporated disaccharide nucleoside can be considered as a chemical (non-enzymatic) method for this site-specific cleavage of DNA. As DNA is built up by nucleosides connected *via* phosphodiester bonds between the 3′-position of one nucleoside and the 5′-position of the adjacent nucleoside, a possible way to introduce a potential cleavage site is to insert a chemical group (*i.e.* ribose) between the 3′-position and the 5′-position of two consecutive nucleotides (Fig. 1). This chemical group should not distort the DNA helix too much and should be sensitive to chemical degradation.

As previously reported (Nauwelaerts *et al.*, 2003), suitable building blocks that fulfill these criteria are 3′-*O*-β-D-ribofuranosyl-2′-deoxy nucleotides. After incorporation into an oligonucleotide, the 5′→5* linked phosphodiester function even renders this oligonucleotide stable against restriction enzymes (*Mbo*I and *Dpn*II) at the insertion site.

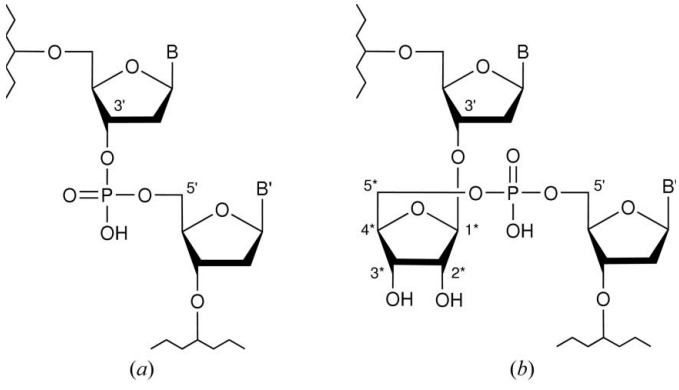

Figure 1
Structure of a natural 3′,5′-phosphodiester linkage (*a*) and of a 5′→5* phosphodiester linkage obtained by inserting a ribose residue between two consecutive nucleotides (*b*).

The cleavage reaction can be obtained using periodate and the ribose fragments removed with NaOH, resulting in the formation of two natural oligonucleotides (one with a free 3′-OH function and the other with a 5′-OPO_3H_2 group), so that both ends are available for a ligation reaction when needed (Nauwelaerts *et al.*, 2003). This means that the complete ribose unit (or the entire alternative chemical group) is removed during the cleavage process. Hence, no nucleotide fragments or genetic information is lost when the cleavage process is carried out.

Therefore, the disaccharide modification might be used as a chemical restriction-site cleaver to insert oligonucleotides into DNA as a sugar-alternative for interference footprinting (Storek *et al.*, 2002) or for the assembly of oligonucleotides with inserted reporter groups.

The already determined NMR structure of the same Dickerson dodecamer with insertion of an extra ribose in the phosphodiester linkage between T*8(20) and C9(21) showed only minor changes in the sugar–phosphate backbone (except from the modified site itself), which was further confirmed by circular dichroism analysis (Nauwelaerts *et al.*, 2003). The linkage between T*8(20) and C9(21) is even able to mimic the connection of two nucleotides as in the non-modified dodecamer (Tjandra *et al.*, 2000).

2. Methods and results

2.1. Synthesis

The modified building block 3′-*O*-*β*-D-ribofuranosyl-dT was synthesized as previously described (Efimtseva *et al.*, 2003). This allows the synthesis of a phosphoramidite building block for oligonucleotide synthesis as reported in Nauwelaerts *et al.* (2003).

2.2. Crystallization

Crystallization conditions of the self-complementary sequence d(CGCGAATT*CGCG), where * indicates the incorporated ribose, were screened using a 24-matrix screen for nucleic acid fragments (Berger *et al.*, 1996) by the hanging-drop vapour-diffusion method at 289 K. As deformed blocks or very thin needle-shaped crystals appeared in five conditions optimization of these conditions had to be performed. Notably, the better needle-shaped crystals appeared only in conditions containing cobalt hexammine. Crystals (Fig. 2) suitable for X-ray diffraction appeared after approximately three weeks using 6 µl droplets of an optimized condition containing 25 m*M* sodium cacodylate buffer pH 5.5, 15 m*M* $MgCl_2$, 22 m*M* NaCl, 2 m*M* cobalt hexammine, 5%(*v*/*v*) 2-methyl-2,4-pentanediol (MPD) and 0.175 m*M* ssDNA equilibrated against 0.5 ml 35%(*v*/*v*) MPD stock solution. The quality of the crystals was tested on a Bruker Smart 6000 CCD system, showing the necessity of using synchrotron radiation.

Figure 2
Typical crystals of the dodecamer d(CGCGAATT*CGCG) (* indicates inserted ribose) obtained by hanging-drop vapour diffusion. The largest crystal is 0.20 mm in length.

Table 1
Data-collection statistics for the dodecamer d(CGCGAATT*CGCG).

Values in parentheses are for the outermost resolution shell.

No. of reflections used	38984
No. of unique reflections	6140 (875)
Resolution range (Å)	20–2.60 (2.74–2.60)
Completeness (%)	96.3 (96.0)
R_{merge} (%)	10.5 (68.1)
Mean $I/\sigma(I)$	16.5 (2.2)
Multiplicity	6.3 (5.9)

2.3. Data collection and processing

A single crystal of dimensions 0.2 × 0.1 × 0.03 mm was used to collect a 96.3% complete data set at EMBL beamline BW7b of the DESY synchrotron in Hamburg. Data were collected on a MAR345 imaging-plate detector with a wavelength of 0.8430 Å, a φ range of 180°, an increment of 1.5° and a crystal-to-detector distance of 435 mm under a liquid-nitrogen cryostream at 100 K. A total of 6140 unique reflections were observed in the resolution range 20–2.6 Å (R_{merge} = 0.105). Although the diffraction pattern showed diffraction to 2.25 Å (Fig. 3), careful analysis showed that the data in the outermost shells was of much lower quality; therefore, a resolution cutoff was chosen at 2.6 Å. Data-collection statistics are given in Table 1.

Data were processed with *MOSFLM* v.6.2.3 (Leslie, 1992) and scaled using *SCALA* v.3.2.5 (Evans, 1997). The latter was used as part of the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

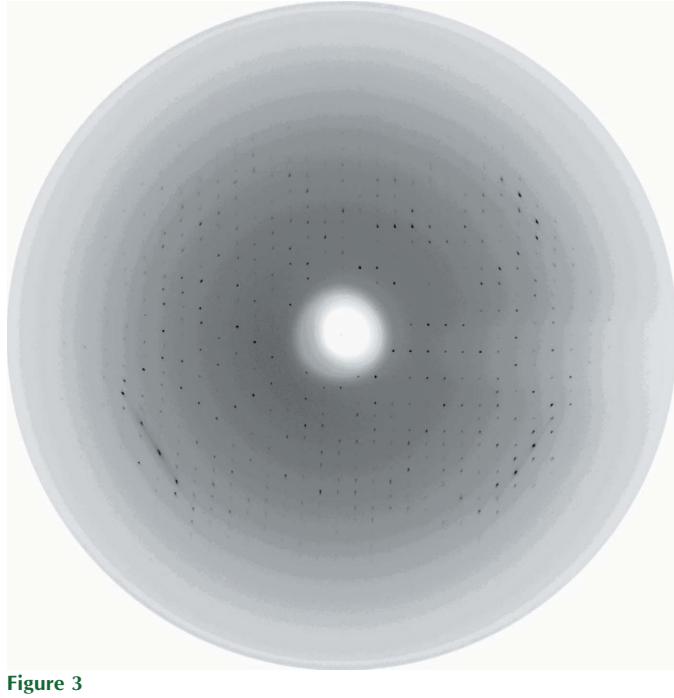

Figure 3
A 1.5° oscillation image of the dodecamer d(CGCGAATT*CGCG) crystal taken on an MAR Research image plate on beamline BW7b (EMBL, Hamburg). The resolution at the edge of the image is 2.25 Å. Base-stacking reflections can be noticed at ∼3 Å.

The crystals belong to the primitive orthorhombic crystal system with unit-cell parameters a = 41.515, b = 57.625, c = 81.389 Å. Inspection of the systematic absences in the (00l) direction indicated that $P222_1$ is the correct space group. When searching the NDB (Berman *et al.*, 1992), only one structure (NDB code PD0007) of a protein–DNA complex (Fraenkel & Pabo, 1998) was found to crystallize in this space group.

The unit cell and space group differ from the Dickerson dodecamer, which suggests a different organization of the helices or even a different helical structure. The Matthews coefficient (V_M) is 3.0 Å^3 Da^{-1} for two double helices in the asymmetrical unit, resulting in a solvent fraction of 58.4%. The volume per base pair of 2028 Å^3 is rather high for B-DNA.

Structure determination by molecular replacement with the program *PHASER* (Storoni *et al.*, 2004) is currently in progress using the Dickerson dodecamer (Berger *et al.*, 1998) as a molecular-replacement model. However, as the unit cell and space-group assignment have already indicated a possibly significantly different structure, SAD phasing is being also performed using the anomalous signal from the P atoms (Ness *et al.*, 2004).

We thank the staff of the EMBL Hamburg Outstation for their support with the synchrotron experiments and A. Van Aerschot for the synthesis of the oligonucleotide. Support from the European Community Research Infrastructure Action under the PF6 'Structuring the European Research Area Programme' contract No. RII3/CT/2004/5060008 is gratefully acknowledged.

References

Berger, I., Kang, C. H., Sinha, N., Wolters, M. & Rich, A. (1996). *Acta Cryst.* D**52**, 465–468.
Berger, I., Tereshko, V., Ikeda, H., Marquez, V. E. & Egli, M. (1998). *Nucleic Acids Res.* **26**, 2473–2480.
Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R. & Schneider, B. (1992). *Biophys. J.* **63**, 751–759.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Efimtseva, E. V., Shelkunova, A. A., Mikhailov, S. N., Nauwelaerts, K., Rozenski, J., Lescrinier, E. & Herdewijn, P. (2003). *Nucleosides Nucleotides Nucleic Acids*, **22**, 359–371.
Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.
Fraenkel, E. & Pabo, C. O. (1998). *Nature Struct. Biol.* **5**, 692–697.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Nauwelaerts, K., Vastmans, K., Froeyen, M., Kempeneers, V., Rozenski, J., Rosemeyer, H., Van Aerschot, A., Busson, R., Lacey, J. C., Efimtseva, E., Mikhailov, S., Lescrinier, E. & Herdewijn, P. (2003). *Nucleic Acids Res.* **31**, 6758–6769.
Ness, S. R., de Graaff, R. A. G., Abrahams, J. P. & Pannu, N. S. (2004). *Structure*, **12**, 1753–1761.
Storek, M. J., Suciu, A. & Verdine, G. L. (2002). *Org. Lett.* **4**, 3867–3869.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Tjandra, N., Tate, S.-I., Ono, A., Kainosho, M. & Bax, A. (2000). *J. Am. Chem. Soc.* **122**, 6190–6196.